UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/610,696	07/02/2003	Amit Srivastava	BBNT-P01-192	1832

28120 7590 09/14/2007

ROPES & GRAY LLP
PATENT DOCKETING 39/41
ONE INTERNATIONAL PLACE
BOSTON, MA 02110-2624

EXAMINER	
SIEDLER, DOROTHY S	
ART UNIT	**PAPER NUMBER**
2626	
MAIL DATE	**DELIVERY MODE**
09/14/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/610,696	SRIVASTAVA ET AL.	
	Examiner	**Art Unit**	
	Dorothy Sarah Siedler	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on 29 June 2007.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-33 is/are pending in the application.

4a) Of the above claim(s) 10,18,23 and 30 is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-9,11-17,19-22,24-29 and 31-33 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 02 July 2003 is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date 6-29-07, 5-25-07

4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

This office action is in response to the amendment filled June 29, 2007. Claims 1-33 are pending; Claims 10,18,23 and 30 are canceled and claims 1, 4, 6, 11, 12, 19, 21, 24, 25, 31, 21 and 33 are amended.

The examiner notes applicant's request for clarification of claim 21. As a typographical error, claim 21 was left out of the 35 U.S.C §103(a) introductory paragraph on page 6 of the office action mailed February 27, 2007. However, the examiner notes that on pages 7-8 of the office action claim 21 was addressed.

Response to Arguments

Applicant's arguments filed June 29, 2007 have been fully considered but they are not persuasive.

Applicant argues that, "***Christensen*** does not disclose or suggest the lexical feature extraction component in claim 1, which is configured to receive text and generate lexical feature vectors relating to the text, the lexical feature vectors including words from the text and syntactic classes of the words" (Remarks page 11). However, the examiner respectfully disagrees. A, "syntactic class" is used by the applicant as, "an indication of the role of a word relative to its surrounding words. For example, possible syntactic classes may indicate whether a word tends to start a sentence, connect phrases together, or end phrases" (Specification page 10). This use of the term "syntactic class" in this instant application is a non-standard usage, since in the art it refers to part of speech. Additionally, ***Christensen*** discloses that, "We formulate the

problem of identifying punctuation marks as that of identifying the last word before a punctuation mark, given a sequence of words and prosodic features. Each word in a text will belong to either one of the punctuation classes or a 'not-preceeding-punctuation' class" (Section 2.2 Linguistic Information, second paragraph). ***Christensen*** determines linguistic features (punctuation marks) based on lexical feature vectors including words and syntactic classes for the words (a sequence of words before the punctuation mark, each word belonging to a punctuation class. The classes include period, comma, and not-preceding punctuation, which are indicative of whether the word starts a sentence, ends a sentence or connects phrases together).

Applicant's arguments with regards to independent claims 12,21 and 25 are similar to those recited in regards to claim 1. Therefore the examiner disagrees with the applicant for similar reasons.

Applicant argues that, "Christensen does not disclose or suggest identifying phrasal boundaries as a linguistic feature. A 'phrasal boundary' as used in the instant application, includes boundaries defined by non-visible information" (Remarks pages 17 and 18). The examiner respectfully disagrees. ***Christensen*** discloses, "The speech recognizer models sentence boundaries and silence using an acoustic model corresponding to pause" (section 2.1 Prosodic Information). ***Christensen*** discloses modeling sentence boundaries (phrasal boundaries) using an acoustic model of a pause (non-visible information).

Claim Rejections - 35 USC § 103

The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Claims 1-9, 11-17, 19-22, 24-29, and 31-33 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Christensen*** ("Punctuation Annotation using Statistical Prosody Models" ISCA Workshop 2001) in view of ***Voutilainen*** ("A Syntax-based part-of-speech analyzer" EACL 1995).

As per claims 1,12,32 and 33, ***Christensen*** disclose a linguistic segmentation tool, method, and device comprising: a lexical feature extraction component configured to receive text and generate lexical feature vectors relating to the text (section 1.1 Prosodic and Linguistic clues to structuring speech, last paragraph and section 2.2 Linguistic Information, *textual clues from the words in the text are used to determine punctuation mark classes*) the lexical feature vectors including words from the text and syntactic classes of the words (section 1.1 Prosodic and Linguistic clues to structuring speech, last paragraph and section 2.2 Linguistic Information, *linguistic features (punctuation marks) are determined based on lexical feature vectors including words and syntactic classes for the word, i.e. a sequence of words before the punctuation mark, each word belonging to a punctuation class. The classes include period, comma, and not-preceding punctuation, which are indicative of whether the word starts a sentence, ends a sentence or connects phrases together*); an acoustic feature extraction component configured to receive an audio version of the text and generate

acoustic feature vectors relating to the audio version of the text (section 1 Introduction, *prosodic features extracted from the audio data are used to determine linguistic features*); and a statistical framework component configured to generate linguistic features associated with the text based on the acoustic feature vectors and the lexical feature vectors (section 2.3 Finite State Model Approach, *the words, punctuation mark classes, and prosodic features are combined into a finite state model*).

Christensen does not disclose wherein the lexical feature extraction component assigns syntactic classes from a set of classes including classes for particular word affixes generated automatically from a corpus of text documents in a given language. ***Voutilainen*** discloses a system that uses a knowledge base for a natural language analyzer (Abstract). ***Voutilainen*** uses a morphological analyzer to assign part of speech, inflectional and derivational tags to input words (classify into classes for particular affixes) (page 158-159, section 2). In addition, ***Voutilainen*** discloses that in a data-driven approach a corpus is used to automatically derive tag information (page 157, section 1, third paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have the lexical feature extraction component assign syntactic classes from a set of classes including classes for particular word affixes in ***Christensen***, since it would enable the system to determine all inflectional and derivational variations of surface forms of input words during lexical analysis.

In addition, it would also have been obvious to one of ordinary skill in the art at the time of the invention to generate classes for affixes automatically from a corpus of

text documents in a given language in ***Christensen***, since one of ordinary skill in the art has good reason to pursue the options within his/her technical grasp to perform a robust and accurate lexical analysis.

As per claim 21, ***Christensen*** discloses a computing device for determining linguistic information for words corresponding to a transcribed version of an audio input stream that includes speech that generates lexical features for the words, including a syntactic class associated with at least one of the words (section 1.1 Prosodic and Linguistic clues to structuring speech, last paragraph and section 2.2 Linguistic Information), generates acoustic features for the audio input stream, the acoustic features being based on at least one of speaker pauses, speaker rate, speaker energy, and speaker pitch (section 1 Introduction, *prosodic features extracted from the audio data are used*), generates the linguistic information based on the lexical features and the acoustic features, and output the generated linguistic information as meta-information embedded in the transcribed version of the audio input stream (section 2.3 Finite State Model Approach and Abstract, *the words, punctuation mark classes, and prosodic features are combined into a finite state model, that information then included as linguistic meta-data for spoken language*). ***Christensen*** does not explicitly disclose the computing device comprising: a processor; and a computer memory coupled to the processor and containing programming instructions that when executed by the processor, cause the processor to perform the previous steps, and wherein the syntactic class being selected

is from a set of classes including classes for particular word affixes generated automatically from a corpus of text documents in a given language.

However, ***Christensen*** discloses that punctuation annotation systems are often used in conjunction with automatic speech recognition systems (1 Introduction), which are typically performed on a computer with a processor and memory containing software instructions. In addition, ***Voutilainen*** discloses a system that uses a knowledge base for a natural language analyzer (Abstract). ***Voutilainen*** uses a morphological analyzer assign part of speech, inflectional and derivational tags to input words (classify into classes for particular affixes) (page 158-159, section 2). In addition, ***Voutilainen*** discloses that in a data-driven approach a corpus is used to automatically derive tag information (page 157, section 1, third paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a computing device with a processor and computer memory in ***Christensen***, since a computer can perform computations from program instructions at a speed far greater than a human can manually, therefore saving processing time.

It would also have been obvious to one of ordinary skill in the art at the time of the invention to have the lexical feature extraction component assign syntactic classes from a set of classes including classes for particular word affixes in ***Christensen***, since it would enable the system to determine all inflectional and derivational variations of surface forms of input words during lexical analysis.

In addition, it would also have been obvious to one of ordinary skill in the art at the time of the invention to generate classes for affixes automatically from a corpus of

text documents in a given language in ***Christensen***, since one of ordinary skill in the art has good reason to pursue the options within his/her technical grasp to perform a robust and accurate lexical analysis.

As per claim 25, ***Christensen*** disclose a method for associating meta-information with a document transcribed from speech, the method comprising: building a language model based on lexical feature vectors extracted from the document, the lexical feature vectors including words and syntactic classifications of the words (section 1.1 Prosodic and Linguistic clues to structuring speech, last paragraph and section 2.2 Linguistic Information, *linguistic features (punctuation marks) are determined based on lexical feature vectors including words and syntactic classes for the word, i.e. a sequence of words before the punctuation mark, each word belonging to a punctuation class. The classes include period, comma, and not-preceding punctuation, which are indicative of whether the word starts a sentence, ends a sentence or connects phrases together*); building an acoustic model based on acoustic feature vectors extracted from the speech (section 1 Introduction, *prosodic features extracted from the audio data are used*); and combining outputs of the language model and the acoustic model in a statistical framework that estimates a probability for associating the meta-information with the document (Abstract and section 2.3 Finite State Model Approach, *the words, punctuation mark classes, and prosodic features are combined into a finite state model to determine linguistic meta-data*).

Christensen does not disclose wherein the syntactic classifications being selected are from a set of classes including classes for particular word affixes generated automatically from a corpus of text documents in a given language. ***Voutilainen*** discloses a system that uses a knowledge base for a natural language analyzer (Abstract). ***Voutilainen*** uses a morphological analyzer assign part of speech, inflectional and derivational tags to input words (classify into classes for particular affixes) (page 158-159, section 2). In addition, ***Voutilainen*** discloses that in a data-driven approach a corpus is used to automatically derive tag information (page 157, section 1, third paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have the lexical feature extraction component assign syntactic classes from a set of classes including classes for particular word affixes in ***Christensen***, since it would enable the system to determine all inflectional and derivational variations of surface forms of input words during lexical analysis.

In addition, it would also have been obvious to one of ordinary skill in the art at the time of the invention to generate classes for affixes automatically from a corpus of text documents in a given language in ***Christensen***, since one of ordinary skill in the art has good reason to pursue the options within his/her technical grasp to perform a robust and accurate lexical analysis.

As per claims 2,20 and 27, ***Christensen*** discloses the linguistic segmentation tool and method of claims 1,12 and 25, wherein the linguistic features include periods, commas

and phrasal boundaries (section 2.2 Linguistic Information). ***Christensen*** does not explicitly disclose the linguistic features including quotation marks and exclamation marks. However, ***Christensen*** does disclose that prosodic and linguistic information combined is used to affectively disambiguate punctuation information in speech (section 1.1 Prosodic and Linguistic clues to Structuring Speech), quotation marks and exclamation marks being common punctuation marks.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to include quotation and exclamation marks as linguistic features in ***Christensen***, since it would enable a system to correctly transcribe text from a spoken utterance, making the transcript usable for other systems such as information retrieval or speech and natural language understanding.

As per claims 3 and 13, ***Christensen*** discloses the linguistic segmentation tool and method of claims 1 and 12, further comprising: a transcription component configured to generate the text based on the audio version of the text (section 1 Introduction, *the ASR system transforms audio into word transcripts*).

As per claim 4, ***Christensen*** discloses the linguistic segmentation tool of claim 1, wherein the statistical framework component includes: an acoustic model configured to estimate a probability of an occurrence of the linguistic features based on the acoustic feature vectors (section 2.3, *prosodic features are combined with punctuation classes into a finite state model to determine punctuation*).

As per claim 5, ***Christensen*** discloses the linguistic segmentation tool of claim 4, wherein the statistical framework component includes: a language model configured to estimate a probability that one of the lexical feature vectors corresponds to a text boundary (section 2.2 Linguistic information, *words and their corresponding punctuation classes are determined, these classes indicative of commas, periods, questions marks etc. which separate text, specifically sentences and words*).

As per claims 6 and 16, ***Christensen*** discloses the linguistic segmentation tool and method of claims 5 and 15, but does not explicitly disclose wherein the statistical framework includes: a maximum likelihood estimator configured to generate the linguistic features based on the probabilities generated by the acoustic model and the language model. However, using a maximum likelihood estimator is well known in the art, by applicant's own admission (specification page 14).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a maximum likelihood estimator in ***Christensen***, since it is a reliable method to combine the acoustic and lexical (punctuation classes) features, and one of ordinary skill has good reason to pursue the known options within his/her technical grasp.

As per claims 7,9,11,17,19,22,24,29 and 31, ***Christensen*** discloses the linguistic segmentation tool and method of claims 1,12,21 and 25, but does not explicitly disclose

wherein the lexical feature vectors additionally include an identification of a structured speech member of the word, wherein the syntactic classes are indicative of a role of the word in the text, wherein the set of classes additionally include syntactic classes defined based on frequently occurring words. ***Voutilainen*** discloses a system that uses a knowledge base for a natural language analyzer (Abstract). ***Voutilainen*** uses a morphological analyzer to assign part of speech, inflectional and derivational tags to input words (page 158-159, section 2). The tags indicate part of speech and inflectional variations, i.e. affix variations, thus indicating the structured speech member as well as its role within the sentence. In addition, ***Voutilainen*** discloses that in a data-driven approach, frequency-based information it automatically derived from a corpus (page 157, section 1, third paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have the lexical feature vectors additionally include an identification of a structured speech member of the word, wherein the syntactic classes are indicative of a role of the word in the text in ***Christensen***, since it would enable the system to determine the linguistic structure of the text, which enables higher-level analysis such as grammatical function assignment and recognizing phrases or other patterns within the text.

In addition, it would also have been obvious to one of ordinary skill in the art at the time of the invention to generate classes for affixes automatically from a corpus of text documents in a given language in ***Christensen***, since one of ordinary skill in the art has good reason to pursue the options within his/her technical grasp.

As per claims 8 and 28, ***Christensen*** discloses the linguistic segmentation tool and method of claims 4 and 25, wherein the acoustic feature vectors are based on prosodic features including at least one of pause, rate, energy, and pitch (section 2.1 Prosodic Information).

As per claim 14, ***Christensen*** discloses the method of claim 12, further comprising: creating a language model configured to estimate a probability that the lexical features correspond to a word boundary based on the lexical features (section 2.2 Linguistic information, *words and their corresponding punctuation classes are determined, these classes indicative of commas, periods, questions marks etc. which separate text, specifically sentences and words*).

As per claim 15, ***Christensen*** discloses the method of claim 14, further comprising: creating an acoustic model configured to estimate a probability of an occurrence of the linguistic information based on the acoustic features (section 2.3, *prosodic features are combined with punctuation classes into a finite state model to determine punctuation*).

As per claim 26, ***Christensen*** discloses the method of claim 25, wherein the meta-information relates to linguistic features of the document (Abstract, *linguistic meta-data*).

Conclusion

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Please see the PTO-892 form.

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dorothy Sarah Siedler whose telephone number is 571-270-1067. The examiner can normally be reached on Mon-Thur 9:30am-5:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone

number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DSS

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER